GeoResJ 6 (2015) 195–201

Contents lists available at ScienceDirect

GeoResJ

journal homepage: www.elsevier.com/locate/GRJ

Bringing dark data into the light: A case study of the recovery of Northwestern Atlantic zooplankton data collected in the 1970s and 1980s

Peter H. Wiebe *, M. Dickson Allison

Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA

ARTICLE INFO

Article history:
Received 10 October 2014
Revised 4 March 2015
Accepted 6 March 2015

Keywords:
Data rescue
Zooplankton biomass
Zooplankton species abundance
Dark data
North Atlantic Gulf Stream Rings

ABSTRACT

Data generated as a result of publicly funded research in the USA and other countries are now required to be available in public data repositories. However, many scientific data over the past 50+ years were collected at a time when the technology for curation, storage, and dissemination were primitive or non-existent and consequently many of these datasets are *not* available publicly. These so-called "dark data" sets are essential to the understanding of how the ocean has changed chemically and biologically in response to the documented shifts in temperature and salinity (aka climate change). An effort is underway to bring into the light, dark data about zooplankton collected in the 1970s and 1980s as part of the cold-core and warm-core rings multidisciplinary programs and other related projects. Zooplankton biomass and euphausiid species abundance from 306 tows and related environmental data including many depth specific tows taken on 34 research cruises in the Northwest Atlantic are online and accessible from the Biological and Chemical Oceanography Data Management Office (BCO-DMO).

1. Introduction

Recent changes in National Science Foundation (NSF) and other agency data policies (NSF11060 [1]; Office of Science & Technology Policy (OSTP) memo 2013 [2]) mandating timely and open access to data and information generated in the course of US funded research have resulted in a relatively rapid change in the culture of data sharing. Technological advances, policy changes, and increased awareness of the need for and benefits of well-curated data make it much more likely that recently generated research results will be made publicly available and in a timely manner (Wallis et al. [3]; Hook et al. [4]). However, many scientific data were generated at a time when the technology for curation, storage, and dissemination were primitive or non-existent, and data sharing was not viewed as essential. In addition, many of the datasets were collected and stored by individuals as small projects that make up the "long tail" of the science enterprise (Heidorn [5]). These smaller projects, in contrast to large projects that involve many investigators, form the bulk of the projects funded by agencies such as NSF. Data from these projects, large and small, have in the past been poorly curated and thus less visible to other scientists, largely not publicly available online, and hence named "Dark Data" (Heidorn [5]). But as Sinha et al. [6] emphasize, without access to the types of historical observations or legacy data that make up the "dark data" in the "long tail" of science, emerging scientific challenges will not be addressable. "...making these data available on demand must be one of the highest priorities for any enterprise seeking to develop a cyberinfrastructure capable of promoting new ways to examine the earth system through time" (Sinha et al. [6]). One international project designed to rescue historical oceanographic data was the IOC/IODE GODAR project, which focused mainly on physical data (Conkright et al. [7]; Caldwell [8]). More recently, the paucity of marine ecosystem data available to conduct cutting edge research and the critical need for the rescue of past data were also highlighted in a recent EarthCube End-User Domain Workshop Report "Articulating Cyberinfrastructure Needs of the Ocean Ecosystem Dynamics Community" (Kinkade et al. [9]) and by Banse [10].

There are significant dark datasets currently unavailable from multidisciplinary programs funded in the 1970s and 1980s such as those that studied the Northwest Atlantic cold-core and warm-core rings (The Ring Group [11]; Joyce and Wiebe [12]).

The Cold-Core Rings (CCR) studies took place between 1972 and 1977, and the Warm-Core Rings (WCR) Program occurred in 1981 and 1982. Large oceanic eddies or rings form when Gulf Stream

* Corresponding author. Tel.: +1 508 289 2313; fax: +1 508 457 2169.
E-mail address: pwiebe@whoi.edu (P.H. Wiebe).

http://dx.doi.org/10.1016/j.grj.2015.03.001
2214-2428/

Table 1
Metadata being sought in the zooplankton data rescue effort. Modified from Anon [38], Annex 3).

Metadata type	Metadata sub-category descriptions
Cruise metadata	• Name of the ship • Investigator-designated Cruise Identifier • Associated Project • Associated Institute • Principal investigator(s) for cruise • Other responsible investigators, and their variable(s) • Cruise or data report
Station metadata	• Station latitude and longitude • Station Month, Day, Year • Station Time (designated as "local", "GMT/UTC", "ship", etc) • Investigator-designated Station Identifier • Meteorological Observations (atmospheric conditions, sea state) • Station Sounding (bottom depth) • Information about any other supplementary/complementary data collected at the same time (same station)
Sampling gear metadata	• Describe the sampling gear used, providing a literature reference if available • If using a "standard" net (e.g., a NORPAC net) was used, be sure to note any modifications to this net • What net mesh size was used (usually in microns) • What was the net opening shape (square or circular) and the opening mouth area or diameter • Was a flowmeter used? When and how was it calibrated?
Net tow metadata	• Towing Method (horizontal, vertical, oblique) • Towing depth-range (a range of starting and ending depths for each net or bottle), or the wire angle and wire out during the tow • Towing Duration (minutes or hours) • Towing Distance (in meters) • Average Towing Speed (knots or meters per second)
Sample processing metadata	• What volume of water was filtered to yield this sample • How were samples preserved, and in what (e.g., 5% buffered formalin) • How were samples processed (summarize the counting, weight, or volume method)? • Was the sampled split (via Folsom splitter or other method)? What was the size of the final aliquot? • Were large plankton removed prior to making biomass measurements? Was a size or volume criteria used in deciding what to remove and what could remain? • Investigator-designated tow, net, or sample identifier
Sample metadata	• Provide the units for each measurement (e.g., #/liter, #/m^3, #/m^2, mg/m^3, mg/haul, ...) • If taxonomic codes, symbols, or abbreviations are used in the data, provide a translation table to help reduce possible misunderstandings of the taxa (e.g., "CfcV" = "Calanus finmarchius copepodite V", ...) • Is an estimate of final uncertainty of the data known?

waters first meander, then separate, forming a ring of Gulf Stream water around a core of cold Slope Water or a core of warm Sargasso Sea water. The CCRs move south or southwest from their point of origin into the Sargasso Sea and are initially 150–300 km in diameter and 2500–3500 m deep. They can persist as identifiable features for up to 2 years. WCRs move to the west/southwest in the Slope Water north of the Gulf Stream. They are 100–200 km in diameter, extend to at least 1500 m deep, and exist for a shorter period of time (usually less than a year) before gradually breaking up and rejoining the Gulf Stream. Both of these kinds of rings form about 5 to 8 times a year.

Rings are particularly interesting to the biologist because species living north and south of the Gulf Stream are distinctly different (Wiebe et al. [13]; Wiebe et al. [14]). Arctic boreal and temperate species from the Slope Water or tropical–subtropical species from the Sargasso Sea are isolated during ring formation within their particular ring structure. Thus, a community of animals from one area is expatriated in the territory of another community of animals. As a ring decays, the water gradually takes on the physical and chemical characteristics of the surrounding non-ring water. Species outside the ring invade the ring habitat while those expatriated go to local extinction (Wiebe and Flierl [15]). This phenomenon provides for a large-scale natural ecological experiment that was the focus of the rings studies.

Data collected during the 1970s in the CCR program were managed by each individual PI separately. For processing and plotting, the data were put onto punch cards and processed by main frame computers such as the Honeywell Sigma 7. Collaborators would meet face to face to discuss the scientific results and share data in the form of written data reports. In the 1980s, the WCR program had a program service office and began to provide some data management services. Most investigators were using microcomputers (manufactured by Commodore, Apple, IBM, and others) and some data were stored on floppy disks. Collaborations between the investigators were conducted at week-long workshops (Wiebe [16]). Some, but not all of the investigators' data and information were stored on a Digital Equipment Corporation minicomputer (VAX 11/780), but when that computer was phased out ~1995, the data were stored on 9-track tapes and they subsequently disappeared. Some of the WCR zooplankton data were summarized in a technical report (Barber and Wiebe [17]). The CTD physical data from many of the cruises were submitted to NODC, but locating the data from these programs is quite difficult without an in-depth knowledge of the program's deployments, etc.

The objective of this paper is to describe the efforts to recover the zooplankton biomass and euphausiid species counts and related environmental data from 34 cruises to the Northwest Atlantic Ocean that were locked in notebooks and old digital file formats, and deposit them into a modern publically available data repository (e.g. the Biological and Chemical Oceanography Data Management Office – BCO-DMO).

1.1. BCO-DMO repository

The Biological and Chemical Oceanography Data Management Office (BCO-DMO) was created and funded by the National Science Foundation (NSF) in 2006 to serve investigators funded by the NSF Biological and Chemical Oceanography Sections to support the scientific research community through improved access to marine biogeochemical and ecological data and information (Anonymous, 2013 [18]). BCO-DMO provides research scientists and others with the systems necessary to work with data from heterogeneous sources with increased efficacy. The BCO-DMO data management system is composed of a metadata database, the distributed client–server JGOFS/GLOBEC data system (Flierl et al. [19]; Glover [20]; Wiebe et al. [21]), and a Web browser with text-based and map-based user interfaces accessing the information and data available from the repository. The metadata database is implemented using the Drupal content management system. These metadata provide the means to discover, access, and reuse data managed by BCO-DMO. The JGOFS/GLOBEC data system provides the means to manage and retrieve the actual data, and any standard Web browser can access the metadata and data. BCO-DMO is a repository for managing data on short- and medium-term time scales; data are routinely submitted to the appropriate national archive.

BCO-DMO uses established controlled vocabularies and ontologies that enable data interoperability, advanced search and discovery (Leadbetter et al. [22]), and the linking of existing data

Fig. 1. Some of the many notebooks where the zooplankton metadata and data may be found.

Table 2
A list of the 34 oceanographic research cruises in the Northwestern Atlantic Ocean between 1972 and 1986 on which zooplankton net tows were taken.

Ship name	Cruise number	Year	Month/day
R/V Gosnold	191	1972	June 7–16
R/V Gosnold	197	1972	August 9–18
R/V Atlantis II	071	1972	September 20/October 14
R/V Chain	111	1973	February 6–19
R/V Knorr	035	1973	November 22/December 3
R/V Knorr	038	1974	March 22/April 4
R/V Atlantis II	085	1974	October 12–23
R/V Chain	125	1975	July 30/August 18
R/V Knorr	053	1975	November 14/December 2
R/V Oceanus	007	1976	June 7
R/V Knorr	062	1976	November 30/December 21
R/V Knorr	065	1977	April 9/May 1
R/V Endeavor	011	1977	July 23/August 18
R/V Knorr	071	1977	October 22/November 18
R/V Atlantis II	101	1978	June 23/July 13
F/V Super Horse	001	1981	April 29–30
F/V Super Horse	002	1981	May 12–13
R/V Oceanus	098	1981	May 29–30
F/V Super Horse	003	1981	June 29
F/V Super Horse	004	1981	July 28–29
F/V Super Horse	005	1981	August 19
R/V Atlantis II	110	1981	September 17/October 7
R/V Oceanus	106	1981	October 27–28
R/V Oceanus	109	1981	November 23–24
R/V Oceanus	111	1981	December 11–15
R/V Oceanus	112	1982	January 5–6
R/V Oceanus	114	1982	February 8–9
R/V Oceanus	116	1982	March 11–16
R/V Oceanus	118	1982	April 18/May 3
R/V Oceanus	121	1982	June 14/July 1
R/V Oceanus	125	1982	August 6–23
R/V Knorr	098	1982	September 24/October 17
R/V Knorr	122	1986	May 28/June 1
R/V Gloria Michelle	–	1986	June 5–9

repositories and networks (Alexander, [23]). They are matched and mapped at three levels: (1) local vocabulary terms, familiar to the originating investigator; (2) intermediate, consistent terms managed by repository custodians (e.g. BCO-DMO); and (3) closest match terms (e.g. from SeaDataNet served by the National Environment Research Council (NERC) Vocabulary Server (NVS) at the British Oceanographic Data Centre (BODC – (http://www.bodc.ac.uk/products/web_services/vocab/)) and shared by the larger community. The standard vocabulary for plankton nets found there and used in this paper is based on the review of net systems that have been used since the late 1880s (Wiebe and Benfield [24]). Multi-level matching and mapping enables retention of important information while improving interoperability of data systems.

2. Methods

Recovering these data started with amassing metadata: how, when and where the zooplankton data were collected. The metadata being sought are summarized in Table 1. As noted above, the data reside in notebooks (Fig. 1), cruise reports, old computer files, and technical reports. However, the crucial element that makes the effort possible is the input of the scientist who conducted the research for which the samples were collected and who remembers many important details about where to look and what to look for. At one time some of the data were entered into a main-frame based database system, which has since disappeared (Hunt and Wiebe [25]).

The search began systematically with the listing of all of the cruises that were part of two rings projects in the 1970s and 1980s, and related sampling programs (Table 2), and then seeking out the information about the zooplankton net tows. These metadata included ship names, station information, tow information, net descriptions rudimentary or otherwise, latitudes and longitudes, times and instrument depths, often including multiple sampling depths with the same net system. The principal focus in the beginning was on recovering the zooplankton biomass data and the counts of euphausiid species from the zooplankton net tows and the associated metadata. Information in the analysis notebooks was not often complete and this required going back into the original cruise log books and crosschecking with other published papers. For some cruises there was a personal log with information that filled in the blanks. There were also errors. The most potentially damaging errors were those of station position. Degrees and decimal minutes were sometimes converted to decimal degrees by simply moving the decimal place and not first dividing the minutes by 60. In addition, sometimes a discrepancy was found between the same information in two different sources. Those errors had to be tracked down using as many other sources as possible.

The original zooplankton biomass and euphausiid catch data and the associated environmental information were acquired using ring nets, net systems, and CTD/rosettes (Conductivity–Temperature–Depth) and Niskin bottles. Ring net, "Bongo" net (McGowan and Brown [26]), and MOCNESS (Multiple Opening and Closing Nets and Environmental Sensing Systems – Wiebe et al. [27,28]) tows were quantified using flow meters calibrated to provide volumes of water filtered for each tow (Fig. 2). The concurrent CTD data from rosettes and the sensors attached to the MOCNESS provided environmental data – temperature, conductivity, and depth.

As the information was assembled, it was typed into a spreadsheet. The decision was made in the beginning of this endeavor to record the data in columns – with one column for each type of data, including species data when they were found. If a different

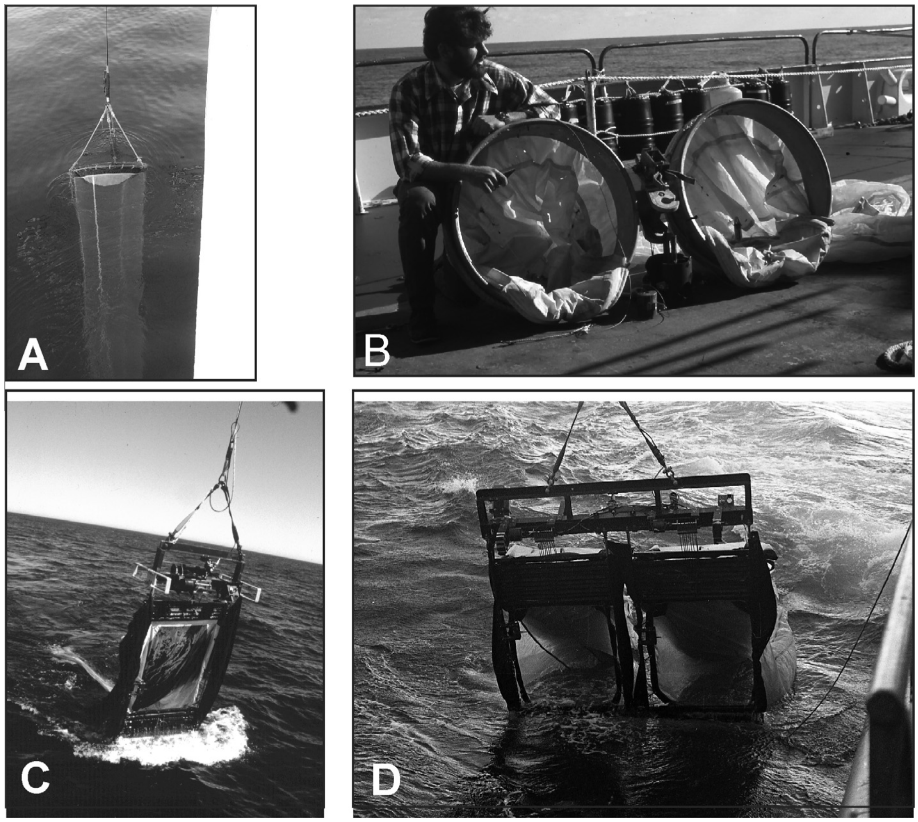

Fig. 2. The net systems used to collect zooplankton in the Northwest Atlantic Ocean between 1972 and 1986. (A) A 1-m diameter ring net. (B) The opening/closing Bongo net system (McGowan and Brown, [26]). (C) A 1-m^2 multiple opening/closing net and environmental sampling system (MOCNESS – Wiebe et al. [27]). (D) A double 1-m^2 MOCNESS (Wiebe et al. [28]).

organization becomes advantageous in the future, the change can be made programmatically.

As the cruise data were collected and digitized, sometimes records of the number of animals counted were found in the same place. Far more often, the sorting and counting of the animals was undertaken later, recorded on data sheets, and stored in separate notebooks. When found, these raw numbers of animals caught were inserted into the spreadsheet with the appropriate line or lines of metadata (e.g. the fraction of the sample sorted and counted). More importantly to this data rescue effort, in a few cases records of tows were found indicating animals were caught, but the data about the catch itself – numbers and kinds of animals – were not found. Equally important, volumes of papers and reports were written using the analyses of catch data and the original data that went into these papers and reports were not readily available. They were not always with the notebooks the cruise metadata came from. More digging into logs and reports, and the originally submitted manuscripts was required. The animals were counted and recorded in a systematic way somewhere and those data needed to be found.

At the same time, in order to make the biomass (e.g. ml/1000 m^3) and raw animal counts into useful data – abundance (numbers/1000 m^3 or /m^2), supporting information was required and became part of the search. The amount of water being filtered to give the number of animals caught is essential to determine abundance. This information comes from the flow meter counts mentioned earlier and then converted to volume of water filtered through the nets. When the flow meters calibration factors were not readily apparent, that information also had to be uncovered. When found, that information was added to the spreadsheet. Biovolumes of animals collected in the nets were usually recorded in the notebooks and, when found, were added to the spreadsheet.

Environmental data (conductivity, temperature, depth – CTD) associated with each net tow were also assembled. In the early 1970s, these data were usually collected on separate hydrographic casts, often by Nansen bottles with reversing thermometers and these data will require a separate recovery effort. With the advent of the MOCNESS in 1974 (Wiebe [27]), CTD data were obtained with electronic sensors (CTD) on the net system. A Neil Brown CTD (Brown, 1974 [29]) was first used as the MOCNESS command and control system, and the environmental data were logged onto 9-track magnetic tapes (Wiebe et al. [27]) during seven cold-core rings cruises and one warm-core ring cruise. This system was replaced with a custom command and control system that used SeaBird conductivity and temperature sensors (Wiebe et al. [28]). In 2003, the 9-track tapes from the CCR cruises were read with a Qualstar tape reader and the binary files saved onto CD-ROMs, and then they were converted from binary to ASCII and transferred to a US GLOBEC Program server. These data were not post-processed until this data recovery project was begun. Processing of the tape data included downloading the individual files from the server into MATLAB format and then removing spikes and other spurious data, and computing salinity using the pressure, temperature, and conductivity values in the Matlab work space. Environmental data from all MOCNESS tow up-traces were saved separately for comparison with the stratified depth intervals sampled while the net system was being hauled back to the surface.

3. Results

Data from 306 tows deployed from 34 research cruises in the Northwest Atlantic, recorded in a spreadsheet (Fig. 3), have been put online at BCO-DMO. Most of the tows resulted in multiple

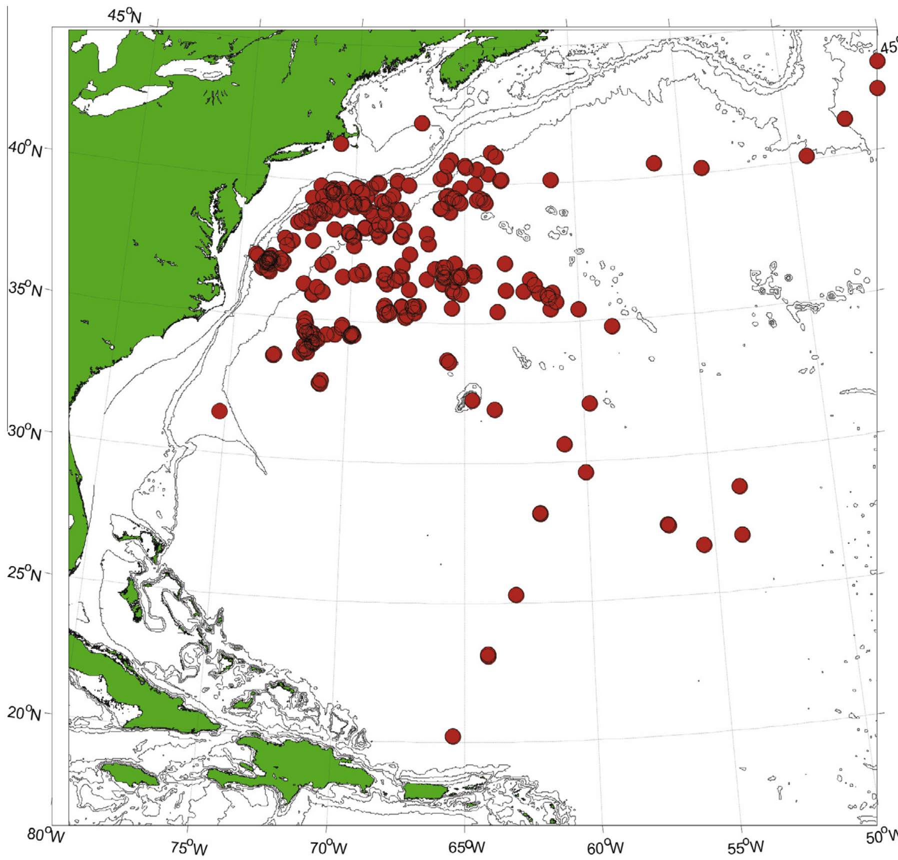

Fig. 3. Locations of the 306 zooplankton net tows taken from 1971 to 1986 on 34 oceanographic research cruises in the Northwestern Atlantic Ocean that are the subject of the data rescue effort.

sampling depths. Thus for the Bongo nets, 2 or 3 strata were sampled; for the 1-m^2 MOCNESS, 8 strata were usually sampled, and for the Double 1-m^2 MOCNESS 16 strata were usually sampled. The number of depth specific samples with biomass and species counts numbered over 2000.

Thirty-four species of euphausiids were identified in the samples from this series of net tows. The raw counts were standardized to numbers per cubic meter using the sample fraction counted and the volume of water filtered by the net. The numbers of individuals under a square meter of sea surface were also determined by multiplying the numbers per m^3 times the depth of water sampled and in the case of the stratified samples summing the values for the individual strata to provide a water column total. A similar computation was done for the biomass data.

Biomass data were originally measured as a displacement volume, and for many tows were converted to carbon units (Wiebe et al. [30]). Equation #1 in Wiebe [31] was used to convert the recovered biovolume data (DV) to carbon (C) using the relationship that relates DV to C and then to milli moles/m^2.

In the Northwest Atlantic the depth of the 10 °C isotherm is diagnostic for the hydrographic province such as the Slope Water (with a shallow isotherm depth) and the Sargasso Sea (with a deep isotherm depth). In Gulf Stream Rings this temperature depth provides a strong indicator of the net tows' hydrographic affinity. In addition, on both CCR and WCR cruises, there was sufficient information to track the movement of the ring center so that a distance from ring center could be calculated. These numbers have been included in the metadata for tows in the vicinity of rings where known. Finally, a site descriptor that provides information about the name of the ring being sampled, and whether sampling took place in the center of the ring or on the edge, or in another region (i.e., Slope Water, Gulf Stream, Sargasso Sea) has also been provided.

For data management purposes, these data have been divided into several distinct groupings: One for cruise metadata, one for total biomass and euphausiid catch data, and one for environmental CTD data. All of these datasets are online and accessible from the Biological and Chemical Oceanography Data Management Office (BCO-DMO) under the project North Atlantic Dark Data (http://www.bco-dmo.org/project/529105).

4. Discussion

This data rescue effort has focused on recovery of the zooplankton biomass and euphausiid species abundance and the associated temperature and salinity data. Other data such as chlorophyll, nutrients, oxygen (CTDs and bottle samples) from the cruises as well as the cruise reports and photographs taken on the cruises remain to be recovered. The final product will be one or more data publications, complete with digital object identifiers (doi).

The data recovery effort described here is just the tip of the iceberg. For some of the tows reported herein there are other groups of zooplankton for which there are counts (copepods, pteropods, chaetognaths) and/or taxa specific abundance and size data that remain to be recovered. In addition, during the Cold-Core and Warm-Core rings programs teams of scientists conducted research in several rings on multidisciplinary ship cruises (Ring Group [11]; Joyce and Wiebe [12]). A large number of instrument systems were used to collect the physical, chemical, and biological data (Fig. 4). The net tow collections for zooplankton were only a part of the ensemble of data resulting from the work at sea. Although some

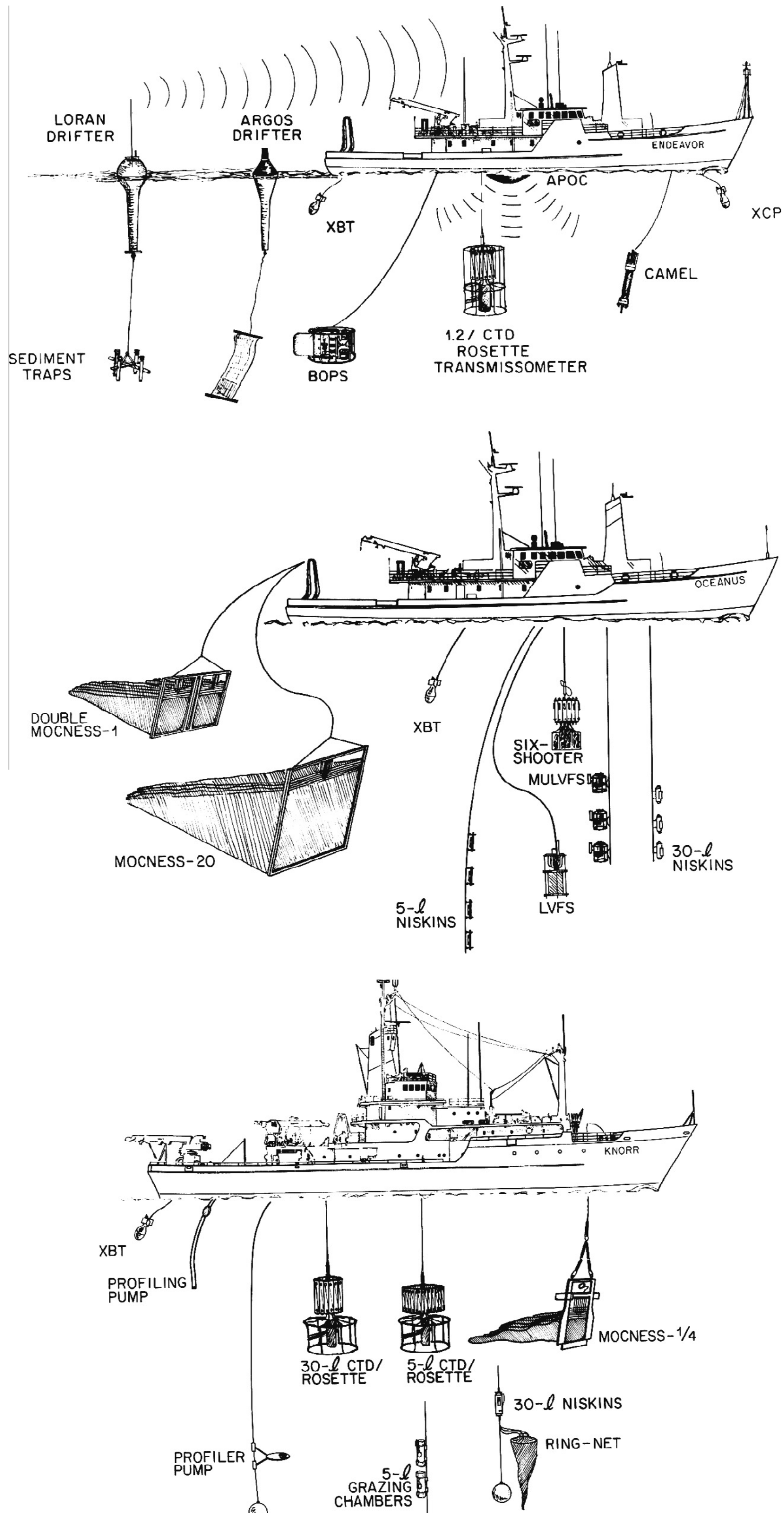

Fig. 4. Over-the-side instrumentation deployed during the multiple ship cruises to study Warm-Core rings (redrawn from Joyce and Wiebe, [12]).

of the physical datasets were submitted to the National Oceanographic Data Center (NODC), most of the chemical and other biological datasets remain in the dark state. Even the data archived at NODC are difficult to find and to relate to these programs since the metadata submitted at the time did not include information about the programs.

Sampling of zooplankton species in the Northwestern Atlantic Ocean between Bermuda and New England Slope Water has had a long, but inconsistent history. Studies of zooplankton of this entire region were conducted between 1937 and 1939 by Clarke [32] and in 1959/1960 by Grice and Hart [33]. Both studies were based on samples taken at stations on the continental shelf, in the Slope Water, and in the Sargasso Sea on a line from Montauk Point, Long Island, NY to Bermuda, but were limited in the depths sampled. A long time-series of zooplankton observations has been carried out in the vicinity of Bermuda by a number of investigators (see review by Steinberg et al. [34]). Surface zooplankton data have been collected between New York and Bermuda with a Continuous Plankton Recorder by NOAA from 1976 to 2013 (Jossi [35]). Other isolated studies have focused on the Gulf Stream as a biogeographical boundary (e.g. Wishner and Allison [36]; Ashjian and Wishner [37]), but are not geographically representative of the Northwest Atlantic region as a whole. The data described in this recovery effort that are now online are from the most area-extensive set of tows that exist in the Northwest Atlantic that have depth specific and environmental information from hydrographically distinct areas.

Banse [10] has clearly articulated the problems facing today's oceanographers who wish to answer questions about how the ocean has changed chemically and biologically in response to the documented shifts in temperature and salinity. Answers to such questions require data collected over the past 50 years or more that are not now available. Data archeology efforts will need to be expanded and support for these activities will need to be enhanced in order to bring into the light the long tail of dark data.

Acknowledgments

We thank Nancy Copley and Robert Groman for assistance in recovering the physical and biological data sets and their assistance in editing this manuscript. We greatly appreciate the helpful editorial comments from three reviewers. This is a contribution from the Biological and Chemical Oceanography Data Management office (BCO-DMO) that is funded by the United States National Science Foundation Grants OCE-1031253 and OCE-1435578.

References

[1] NSF11060. U.S. National Science Foundation. 2011. *Division of Ocean Sciences sample and data policy* (document number nsf11060). Retrieved from <http://www.nsf.gov/pubs/2011/nsf11060/nsf11060.pdf>.

[2] OSTP Memo 2013. U.S. Office of Science and Technology. *Increasing access to the results of federally funded scientific research*. Washington, DC. Retrieved from <http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf>.

[3] Wallis JC, Borgman CL, Mayernik MS, Pepe A. Moving archival practices upstream: an exploration of the life cycle of ecological sensing data in collaborative field research. Int J Digit Curation 2008;3(1):114–26.

[4] Hook LA, Santhana Vannan SK, Beaty TW, Cook RB, Wilson BE. Best practices for preparing environmental data sets to share and archive; 2010. Available online (<http://daac.ornl.gov/PI/BestPractices-2010.pdf>) from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA. doi:10.3334/ORNLDAAC/BestPractices-2010.

[5] Heidorn PB. Shedding light on the dark data in the long tail of science. Libr Trends 2008;57(2):280–99. http://dx.doi.org/10.1353/lib.0.0036.

[6] Sinha AK, Thessen AE, Barnes CG. Geoinformatics: towards an integrative view of Earth as a system. In: Bickford ME, editor. The Web of Geological Sciences: advances, Impacts, and Interactions. Geological Society of America; 2013. http://dx.doi.org/10.1130/2013.2500(19) (GSA Special Paper 500, pp. 591-604).

[7] Conkright ME, Antonov JI, Baranova O, Boyer TP, Garcia HE, Gelfeld R, et al. World ocean database 2001, volume 1: introduction. In: Sydney Levitus, editor. NOAA Atlas NESDIS 42. Washington, DC: U.S. Government Printing Office; 2002 [167 pp.].

[8] Caldwell P. NOAA support for global sea level data rescue. Earth System Monitor. 14 (1): 1–2, 5–6. Silver Spring, MD: NOAA; 2003. Available from U.S. NODC Oceanographic Data Center.

[9] Kinkade D, Chandler C, Glover D, Groman R, Kline D, Nahorniak J, et al. Articulating cyberinfrastructure needs of the ocean ecosystem dynamics community. Earthcube End-User Domain Workshop Report; 2013. Final report submitted to earthcube.org. Summary at <http://workspace.earthcube.org/sites/default/files/files/document-repository/OcanEcosystemDynamicsEndUserWorkshop.pdf>.

[10] Banse KM. Assessing ocean changes without data centers? Front Mar Sci 2014. http://dx.doi.org/10.3389/fmars.2014.00029.

[11] Ring Group. Gulf Stream cold core rings: their physics, chemistry, and biology. Science 1981;212:1091–100 (Backus RH, Flierl GR, Kester D, Olson DB, Richardson PL, Vastano AC, et al.).

[12] Joyce TM, Wiebe PH. Warm core rings of the Gulf Stream. Oceanus 1983;26(2):34–44.

[13] Wiebe PH, Hulburt EM, Carpenter EJ, Jahn AE, Knapp III GP, Boyd SH. Gulf Stream cold core rings: large scale interaction sites for open ocean planktonic communities. Deep-Sea Res 1976;23:695–710.

[14] Wiebe PH, Copley NJ, Boyd SH. Coarse-scale horizontal patchiness and vertical migration in newly formed Gulf Stream warm-core ring 82-H. Deep-Sea Res. 1992;39(Suppl. 1):247–78.

[15] Wiebe PH, Flierl GR. Euphausiid invasion/dispersal in Gulf Stream cold core rings. Aust J Mar Freshwater Res 1983;34(4):625–52.

[16] Wiebe, P. [ed] 1983. Interdisciplinary study of Gulf Stream warm-core rings physics, chemistry, and biology. Report of the warm-core rings data workshop; July 6–19, 1983. Whispering Pines Conference Center, W. Alton Jones Campus, University of Rhode Island, West Greenwich, Rhode Island. 30 Pages. <http://dmoserv3.bco-dmo.org/data_docs/NAtlDarkData/Wiebe_Report_WCR_DataWorkshop_AltonJones_July1983.pdf>.

[17] Barber, V, Wiebe, P. Zooplankton biomass and related MOCNESS data for tows taken on the 1981–82 warm-core rings cruises. Woods Hole Oceanographic Institution (WHOI). Technical report 85-3; 1985.

[18] Anonymous. Description of the biological and chemical oceanography data management office (BCO-DMO); 2013. <http://www.bco-dmo.org/sites/default/files/BCO-DMO_Introduction_v3.pdf>.

[19] Flierl G, Bishop JKB, Glover DM, Paranjpe S. Data management for JGOFS: theory and design. In: Proceedings of ocean climate data workshop. February 18–21, 1992. Goddard Space Flight Center, Greenbelt, MD: US NOAA and NASA; 1992 (pp. 229–249).

[20] Glover DM. Taking care of the legacy: data management in US JGOFS. Oceanography 2001;14(4):106–7.

[21] Wiebe PH, Groman R, Chandler C. Data management in the US Global Ocean Ecosystem Dynamics Program (US GLOBEC). In: The US GLOBEC program: climate change and marine ecosystems special issue of CURRENT; 2011 27 (2): 30-33.

[22] Leadbetter AM, Lowry RK, Clements DO. Putting meaning into NETMAR—the open service network for marine environmental data. Int J Digit Earth 2014;7:811–28.

[23] Alexander, P. 2011 Ontologies: relational vocabularies. In: The MMI guides: navigating the world of marine metadata. <http://marinemetadata.org/guides/vocabs/ont>.

[24] Wiebe PH, Benfield MC. From the Hensen Net toward four-dimensional biological oceanography. Prog. Oceanogr 2003;56(1):7–136.

[25] Hunt M, Wiebe PH. A data base for zooplankton net tow data. W.H.O.I. technical report 80–28; 1980. 65 pp.

[26] McGowan JA, Brown DM. A new opening-closing paired zooplankton net. Scripps Institute of Oceanography Reference; 1966, 66/23. p. 1–56.

[27] Wiebe PH, Burt KH, Boyd SH, Morton AW. A multiple opening/closing net and environmental sensing system for sampling zooplankton. J Mar Res 1976;34:313–26.

[28] Wiebe PH, Morton AW, Bradley AM, Backus RH, Craddock JE, Barber V, et al. New developments in the MOCNESS, an apparatus for sampling zooplankton and micronekton. Mar Biol 1985;87:313–23.

[29] Brown NL. A precision CTD microprofiler. Engineering in the Ocean Environment, Ocean '74 – IEEE international conference on, Vol. 2; 1974. p. 270–278. doi: 10.1109/OCEANS.1974.1161443.

[30] Wiebe PH, Boyd SH, Cox J. Relationships between zooplankton displacement volume, wet weight, dry weight, and carbon. Fish Bull 1975;73(4):777–86.

[31] Wiebe PH. Functional regression equations for zooplankton displacement volume, wet weight, dry weight, and carbon: a correction. Fish Bull 1988;86:833–5.

[32] Clarke GL. Comparative richness of zooplankton in coastal and offshore areas of the Atlantic. Biol Bull 1940;78:226–55.

[33] Grice GD, Hart AD. The abundance, seasonal occurrence and distribution of the epizooplankton between New York and Bermuda. Ecol Monogr 1962;32(4):287–309.

[34] Steinberg DK, Carlson CA, Bates NB, Johnson RJ, Michales AF, Knap AH. Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry. Deep-Sea Res II 2001;48(8–9):1405–47.

[35] Jossi JW. US NOAA Fisheries and UK SAHFOS CPR surveys: comparison of methods and data. J Plankton Res 2010;32:1623–31. http://dx.doi.org/10.1093/plankt/fbq032.

[36] Wishner KF, Allison SK. The distribution and abundance of copepods in relation to the physical structure of the Gulf Stream. Deep-Sea Res 1986;33:705–31.

[37] Ashjian CJ, Wishner KF. Temporal persistence of copepod species groups in the Gulf Stream. Deep-Sea Res I 1993;40:483–516.

[38] Anon. Report of the working group on zooplankton ecology. Gijón, Spain: ICES CM 2003/C:01; 24–26 February 2003. 52 pp.